

# EVIDENCE-BASED INTERVENTION REVIEW CRITERIA & REVIEW RUBRICS

*Applied criteria to identify evidence-based interventions*

# TABLE OF CONTENTS

<b>1. Introduction .....</b>	<b>3</b>
Overview	
Establishing Criteria to Identify Evidence-Based Interventions	
Reviewer Guide and <i>Quality Review Rubric</i>	
<b>2. Appendices .....</b>	<b>7</b>
A: Reviewer Guide to the Criteria for Evidence	
B: <i>Quality Review Rubric</i>	
C: Additional Rubrics	

# 1. Introduction

Overview	<p>This document contains information regarding the review criteria and related rubrics used to identify evidence-based PK-6 literacy interventions. The rubrics were applied to specific interventions that included one or more of the following domains of literacy specific skills: print concepts, phonological awareness, alphabet knowledge, phonics/decoding, irregular/sight words, fluency, oral language, vocabulary, comprehension, spelling and writing. For more information about the reviews, please see <i>Reviewed Intervention Brief 2016</i>.</p>
Establishing Criteria to Identify Evidence-Based Interventions	<p>Criteria were established to determine that a particular intervention showed evidence of effect for a given area of literacy, for a given grade, and for a given tier of support (classwide<sup>1</sup>, targeted<sup>2</sup>, and/or intensive<sup>3</sup>). There were three major action steps in developing the criteria:</p> <ol style="list-style-type: none"><li>1. <b><u>Establish and convene task group.</u></b> The purpose of the task group was to establish criteria, the review rubric, and process to review universal, targeted and intensive programs and interventions. This 12 member group was established in 2014, convened weekly and included one national expert, who collaborated across two additional national experts:<ul style="list-style-type: none"><li>• Dr. Jeanne Wanzek<sup>4</sup> – Florida State University (lead), expertise in K-6 literacy, served as the primary developer of the criteria and rubric,<ul style="list-style-type: none"><li>○ Dr. Laura Justice – Ohio State University, expertise in PK literacy, worked directly with Dr. Wanzek to ensure PK representation,</li><li>○ Dr. Andy Porter – University of Pennsylvania, expertise in alignment of standards, worked to develop alignment criteria and rubric.</li></ul></li></ul><p>The purpose of the national expert was to work directly with the task group to develop the criteria and review rubric based on current research. Dr. Wanzek served as the primary developer of the criteria/rubric with guidance from the task group.</p></li><li>2. <b><u>Establish and convene vetting group.</u></b> The purpose of this group was to provide feedback, guidance and input on products and processes developed by the task group. This 31 member group was established in 2014, and met a total of three times to provide critical feedback and guidance to the work.</li><li>3. <b><u>Develop Criteria and Rubric.</u></b> The task group, including the national expert, were primarily responsible for developing the</li></ol>

<sup>1</sup> This included evidence-based interventions or programs shown to be effective with entire classrooms of students.

<sup>2</sup> This included evidence-based interventions or programs shown to be effective at Tier 2/targeted level of an MTSS framework.

<sup>3</sup> This included evidence-based interventions or programs shown to be effective at Tier 3/intensive level of an MTSS framework.

<sup>4</sup> Dr. Jeanne Wanzek has since moved to Vanderbilt University.

---

criteria and rubric. Dr. Wanzek reviewed extant research, and developed criteria/rubric drafts. The work was reviewed weekly, and guidance provided by the task group. Feedback and input was provided by the vetting group to incorporate directly into products.

In order to determine whether a particular intervention showed evidence of effect for a given area of literacy, for a given grade, and for a given tier of support, Dr. Wanzek developed criteria to determine the extent of evidence of interventions. The extent of evidence was based on the quality, replicability, generalizability and positive results of published research and/or technical reports for any given intervention.

The national expert, working with the task group, adapted criteria from previous work at the national level, including the What Works Clearinghouse (WWC) and the National Center on Intensive Interventions (NCII). For details about this work, see Dr. Jeanne Wanzek's white paper at: [iowareadingresearchcenter.org](http://iowareadingresearchcenter.org)

The adapted criteria is based on internal and external validity as primary considerations of the quality, as well as the extent of evidence within and across studies. Brief descriptions of each follow:

**Internal Validity.** Internal validity addresses how well a research study was designed to reduce the impact of things not being studied. This increases the likelihood that positive results are due to the intervention, and not things outside of the study. Internal validity criteria were: Research Design, Evidence of Confounding Factors, Group/Person Conducting the Study, Developer of Assessment, Data Collection and Adequacy of Measures, and Data Analysis Methods.

- **External Validity.** External validity addresses the extent to which the study and its findings apply to other practical settings beyond the controlled research study. External validity criteria were: Implementation, Reading Domains Addressed, Student Outcomes Measured, and Treatment Acceptability.
- **Overall Findings.** This area addresses the extent of positive overall findings of a given study [RFP Review]. Specifically, the study findings must demonstrate statistical significance or an effect size of .25 or higher on one or more reading domains and no statistically significant negative effects. In the case of single case designs, the reading data must demonstrate at least three instances of an effect.

---

A *Quality Review Rubric* was developed based on the criteria. The criteria descriptions are in Appendix A; the rubric is in Appendix B. The rubric was used to review interventions for both the RFP and Iowa Reviews.

---

## Reviewer Guide and *Quality Review Rubric*

Once the quality criteria were established, Dr. Wanzek, working with the task group, developed:

1. **A Reviewer Guide to the Criteria and Rubric – Appendix A.** This guide provides detailed information about each criteria within the rubric:
  - a ***General Information.*** This includes the following: Intended Grade Levels, Reading Domains Covered, Recommended Dosage of Program, Number of Lessons Available, Placement Assessment Included, Intended Population of Students, Recommended Implementers, Recommended Grouping Formats, Parent/Home Connection Strategies/Materials Included, Number of Studies Submitted, Number of Peer-Reviewed Studies Submitted, Costs of Materials, Training, Hours of Training and any Additional Costs
  - b ***Evidence for Program Effectiveness.*** This includes the following:
    - *Internal Validity*----Study Design, Group/Person Conducting the Study, Developer of Assessment, Technical Adequacy of measures to determine effect size or evidence of improvement, Data Collection, Data Analysis, Evidence of Confounding Factors
    - *External Validity*----Group/Person Implementing Intervention/Program, Dosage: Session Time and Frequency, Fidelity of Implementation, Reading Domains Addressed, Grouping Format, Student Outcomes Measured, Treatment Acceptability
    - *Findings*----Overall Findings and Long Term Findings
    - *Summary of Evidence*----Participants and Extent of Evidence
2. **A Quality Review Rubric – Appendix B.** The rubric has all the criteria in a usable review rating form. Generally, there are three levels of ratings: Desired, Acceptable and Undesired. Scores are established across each area – recommended scores are included within the rubric in Appendix B. The rubric is designed to be used in conjunction with the Reviewer Guide.

Although Appendix A & B provide neither direct technical assistance nor a step-by-step guide for reviewing evidence, both the rubric and guide are available for teams to use to establish the evidence-base of any given intervention.

---

To use the Reviewer Guide and *Quality Review Rubric*, reviewer teams should be established, and follow similar procedures as outlined below:

1. **Determine Interventions to Review.** Within the team, determine how to select interventions to review. This may be done via an online survey to understand current intervention practices, or reviewing current literature, or in consultation with local Area Education Agencies or Institutes of Higher Education.
2. **Determine Study Criteria.** Teams should complete extensive literature reviews for each intervention, and therefore should put in place study-inclusion criteria. See 3. The Reviews – Iowa Review 2015 for sample study criteria.
3. **Implement Tight Review Procedures.** Teams should be trained on the Reviewer Guide and *Quality Review Rubric* before using it to make decisions. Training should include an inter-rater reliability component to ensure reviewers can effectively and consistently apply the criteria. At least two reviewers should be assigned to the same information to review and score independently. All reviewers should be free of any conflicts of interest. Subsequent to final scoring, discrepancies must be reconciled and final scores submitted for final analysis.
4. **Apply Thresholds.** After all reviews are completed across all studies for each intervention, the threshold of acceptable in the area of Extent of Evidence should be used to determine whether any given intervention is evidence-based for a particular area of literacy in a specific grade or age-span. This should include the following criteria: One study with high internal and external validity and positive findings with no studies showing negative findings. In the case of Intensive Interventions where a single case was used, five or more studies with high internal and external validity.

2. Appendices includes the following:

A. Reviewer Guide to the Criteria for Evidence..... 7

B. Quality Review Rubric..... 16

C. Additional Rubrics [RFP Review only]..... 23

A. Reviewer Guide to the Criteria for Evidence	<p>The Reviewer Guide was developed by Dr. Jeanne Wanzek, and is designed to be used with the <i>Quality Review Rubric</i> within a very tight review process [described in 1. <i>Review Guide and Quality Review Rubric</i>].</p> <p>To best use the Reviewer Guide and rubric, a review team must use one or more of the following items:</p> <ol style="list-style-type: none"><li>1. Materials or studies submitted by a vendor in response to a Request for Information or Proposal; or</li><li>2. Materials or studies obtained online or through a journal search; or</li><li>3. The academic intervention tool chart from the reviews conducted by the National Center on Intensive Intervention (NCII) located at <a href="http://www.intensiveintervention.org/chart/instructional-intervention-tools">http://www.intensiveintervention.org/chart/instructional-intervention-tools</a>. Generally, the reviews for NCII are for targeted and/or Intensive interventions and not classwide interventions. Be sure to click on the program link for each specific study to receive the detailed information needed to complete the relevant aspects of the rubric. For any rubric information not available in the NCII review, the study reviewed can be located to provide the additional information.</li></ol>
--	--

Reviewer Guide to the Criteria and Rubric

Author: Dr. Jeanne Wanzek

Instructions

General Information

Use the materials to identify the following information. If using a review in the tools chart from the National Center on Intensive Interventions as one piece of information, click on the program to find descriptive information on the intervention.

Intended Grade Levels

Mark the grade levels the program is intended to serve

Reading Domains

Mark the domains the program is intended to address.

Print Concepts	Recognizing the components of written language (e.g., words, sentences, print moving from left to right, etc.)
Phonological Awareness	Recognizing and manipulating the sound system in spoken language
Alphabet Knowledge	Naming, distinguishing, and writing the letters of the alphabet
Phonics/Decoding	Identifying sound-symbol relationships and using them to read and spell words.
Irregular/Sight Words	Reading words in which some or all of the letters do not represent most commonly associated sounds; Recognizing words (regular and irregular) by sight
Fluency	Ability to perform reading skills with quickly, accurately, and with proper expression
Oral Language	Understanding of the phonology, grammar, morphology, vocabulary, discourse, and pragmatics of the English language
Vocabulary	Understanding and using words when listening, speaking, reading, and writing
Comprehension	Understanding the intended meaning of spoken and written language
Spelling	Ability to write or name the letters of a word
Writing	Composing text to express ideas or opinions
Other	Any other literacy-related domains; please specify

### Recommended Dosage of Program

Identify the number of weeks of instruction provided/recommended in the program, the recommended session length per lesson (in minutes), and the recommended frequency of lessons (number of sessions per week). Mark N/R if the information is not provided.

### Number of Lessons Available

Identify the number of lessons provided in the base materials and the number of lessons provided in any extended/supplemental materials that can be purchased. Mark N/A if base or extended lessons are not provided/available.

### Placement Assessment Included

Indicate whether a placement assessment is included to help teachers begin students at the appropriate lesson to meet their needs.

### Intended Population of Students

Mark the population(s) of students the program is intended to serve. Mark N/R if the information is not reported.

### Recommended Implementers

Mark the recommended implementers for the program. Mark N/R if the information is not reported.

### Recommended Grouping Formats

Mark the grouping formats that are recommended for program implementation. Mark N/R if the information is not reported.



## Parent/Home Connection Strategies/Materials Included

Indicate whether the program includes strategies and/or provides specific materials for connecting with parents or families at home.

## Study Information

Identify the number of primary analysis studies of the intervention that were submitted. Of these studies, identify how many are published in peer-reviewed outlets.

## Costs

Identify the cost per pupil of the base materials, the training cost, the recommended hours of training to implement the program, and any costs for extended materials or additional training. Mark N/A if the cost category does not apply to the program. Mark N/R if the information is not reported.

# Evidence for Program Effectiveness

For each unique, primary analysis study of the program submitted or reviewed previously by the National Center on Intensive Interventions, complete the ratings on internal validity, external validity, and findings. A primary analysis study of the program is an examination of the effect of an intervention on a particular sample (e.g., a set of students or schools). Studies of the way in which an intervention was implemented without evidence of impact, literature reviews, or meta-analyses are not considered primary analyses of the program.

## Internal Validity

### a. Research Design (this information is available in reviews on NCII under “Design”)

*Desired State/Optimal:* Studies that randomly assign participants to study conditions (students or clusters can be randomly assigned) and demonstrate no significant overall or differential attrition (attrition bias). The What Works Clearinghouse (WWC) guidance is used in determining the significance of the overall and differential attrition or attrition bias (p. 11-14):

[http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v3\\_0\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf)

A higher rating is also given for data demonstrating the intervention and comparison groups were not statistically significantly different at pretreatment. In the case of Intensive Interventions where a single case design may be used, the baseline data for all cases should be stable prior to intervention implementation.

*Acceptable:* Randomized control trials with significant overall or differential attrition, or treatment-comparison studies without random assignment (quasi-experimental studies) with study conditions matched on several pretreatment variables (e.g., demographics, reading achievement), referred to as tenable quasi-experiments on the NCII reviews.

For Intensive Interventions, single case design with at least 3 replications and at least 3 data points per phase of the design.

Intervention and comparison groups that are statistically significantly different at pretreatment must be within .25 SD difference at pretreatment, and outcomes of interest must be adjusted for these pretreatment differences in the analyses.

In the case of Intensive Interventions where a single case design is used, the majority of cases (must be more than 3 cases) must show stable baseline data.

*Undesired State:* Study designs not included in desired or acceptable ratings, quasi-experimental designs without adequate matching, or single case designs with less than 3 replications of effects or less than 3 data points per phase of the design.

**b. Group/Person Conducting Study**

*Desired State/Optimal:* The study was conducted by independent evaluators who do not have a conflict of interest related to the intervention's impact (e.g., financial interest in the intervention).

*Acceptable:* The study was conducted by those who have a possible conflict of interest (e.g., developers of the intervention but with no financial interest in the intervention).

*Undesired State:* The study was conducted by persons with a direct conflict of interest in the intervention.

**c. Developer of Assessment**

*Desired State/Optimal:* The majority of the reading domain assessments used in the study were developed by someone other than the authors/developers/vendors of the intervention.

*Acceptable:* Less than half of the reading domain assessments used in the study were developed by the vendor but not the authors/developers of the intervention.

*Undesired State:* All of the reading domain assessments used in the study were developed by the authors/developers of the intervention.

**d. Technical Adequacy of Measures Used to Determine ES or Evidence of Improvement (reliability information is available in reviews on NCII in the measures table under "Measures Targeted" and "Measures Broader")**

*Desired State/Optimal:* Reliability coefficients for all reading domain measures used in the study were  $\geq .70$  with most of the reliability coefficients  $\geq .80$ . Validity coefficients for all reading domain measures used in the study were  $\geq .30$  with most of the validity coefficients  $\geq .50$

*Acceptable:* Reliability coefficients for all reading domain measures used in the study were  $\geq .70$ . Validity coefficients for all reading domain measures used in the study were  $\geq .30$ .

*Undesired State:* Reliability coefficients for most reading domain measures used in the study were  $< .70$  or reliability coefficients were not provided for most reading domain measures. Validity coefficients for most reading domain measures used in the study were  $< .30$  or validity coefficients were not provided for most reading domain measures.

**e. Data Collection**

*Desired State/Optimal:* The reading domain outcome data was collected by assessors blind to the study conditions, and reliability of the assessors  $\geq .90$ . In the case of Intensive Interventions when single case design is used, there is more than one assessor and reliability of the assessors is  $\geq .90$ .

*Acceptable:* The reading domain outcome data was collected by assessors blind to part of the study conditions and reliability of the assessors is  $\geq .80$ .

*Undesired State:* The reading domain outcome data was collected by assessors who were not blind to study conditions, or the reliability of the assessors  $< .80$ . In the case of Intensive Interventions where single case design was used, a single assessor was used.

**f. Data Analyses**

*Desired State/Optimal:* The unit of assignment to condition (e.g., student, class, school) matches the unit of analysis and/or clustering/nesting of students in the unit of assignment is taken into account, OR when multiple comparisons are conducted the p value is adjusted to control Type I error.

*Acceptable:* The unit of assignment to condition (e.g., student, class, school) matches the unit of analysis and/or clustering/nesting of students in the unit of assignment is taken into account, OR when multiple comparisons are conducted the p value is adjusted to control Type I error.

*Undesired State:* Analyses do not match the design of the study (e.g., the unit of assignment was school and only student data are modeled)

**g. Evidence of Confounding Factors**

Confounding Factors include any indicator that the study conditions differed in ways beyond the intervention of interest or that may affect the findings in ways that are not related to the impact of the intervention. For example, when the data of only students who responded to the intervention are analyzed the findings related to the impact of the intervention may be inflated. Another example, is the outcome measure being collected in different manner for the treatment and comparison conditions (e.g., different time, different personnel, different outcome measures). The What Works Clearinghouse provides several other examples of confounding factors:

*The most common type of confounding occurs when there is one unit (e.g., a single teacher, classroom, school, or district) in one or both of the conditions. When only one unit is used in one of the conditions, there is no way to distinguish between the effect of the intervention and that unit (e.g., a particularly good teacher). For example, if all of Mrs. Smith's students use one reading program while all of Mr. Jones's use another (even if students were randomly assigned to the class/condition), differences between the outcomes of students in the two conditions may be due to the program, the teacher, or both. Another example of confounding occurs when the characteristics of the units in each group differ systematically in ways that are associated with the outcomes. For example, a small group of teachers in a master's program implements the intervention, while the comparison group is made up of teachers with bachelor's degrees.*

While the above provide several examples of confounding factors, reviewers should consider any information in study that indicates the study condition differ in ways beyond the intervention of interest or that may affect the findings in ways that are not related to the impact of the intervention.

*Desired State/Optimal:* No evidence of confounding factors that could have interfered with results.

*Undesired State:* Evidence of confounding factors that could have interfered with results

## **External Validity**

**a. Group/Person Implementing Intervention**

*Desired State/Optimal:* The intervention was implemented in a school setting by school personnel.

*Acceptable:* The intervention was implemented in a school setting with specially trained personnel outside of school personnel (e.g., researcher).

*Undesired State:* The intervention was implemented outside of a school setting such as a clinic by specially trained personnel.

**b. Dosage: Session Time and Frequency (this information is available in reviews on NCII under "Duration of Intervention")**

For this category use the information on the general information page to determine the recommended session length per lesson, and the recommended frequency of lessons (number of sessions per week). Then rate the study implementation relative to these recommended dosage amounts.

*Desired State/Optimal:* Session time and frequency in the study were aligned with recommendations for the intervention.

*Undesired State:* Session time and frequency in the study were not aligned with recommendations for the intervention.

**c. Fidelity of Implementation (this information is available in reviews on NCII under “Fidelity of Implementation”)**

*Desired State/Optimal:* Detailed data are provided that the intervention was implemented as intended during the study, the data demonstrate there was generally high implementation (75% or higher), and intercoder agreement for the collection of fidelity data is  $\geq .90$ .

*Acceptable:* Limited data are provided that the intervention was implemented as intended during the study and the data demonstrate there was generally high implementation (75% or higher), or intercoder agreement for the collection of fidelity data is  $\geq .80$ .

*Undesired State:* No data related to the implementation of the intervention during the study are provided, the intervention was not implemented as intended (less than 75%), or intercoder agreement for the collection of fidelity data  $< .80$ .

**d. Reading Domains Addressed**

This category is rated only for Targeted Interventions.

*Desired State/Optimal:* Three or more foundational reading domains (e.g., phonological awareness, phonics, fluency, vocabulary, comprehension) are addressed in the intervention.

*Acceptable:* At least two foundational reading domains (e.g., phonological awareness, phonics, fluency, vocabulary, comprehension) are addressed in the intervention.

*Undesired State:* Fewer than two foundational domains (e.g., phonological awareness, phonics, fluency, vocabulary, comprehension) are addressed in the intervention.

**e. Grouping Format (this information is available in reviews on NCII under “Administration Group Size”)**

For this category use the information on the general information page to determine the recommended grouping format. Then rate the study implementation relative to the recommended grouping format.

*Desired State/Optimal:* The grouping format(s) implemented in the study aligned with the recommended grouping for intervention.

*Undesired State:* The grouping format(s) implemented in the study did not align with the recommended grouping for intervention.

**g. Student Outcomes Measured (this information is available in reviews on NCII under “Measures Targeted” and “Measures Broader”)**

We use the descriptions of measures from the National Center on Intensive Interventions to consider the outcome measures used in the study.

*Targeted measures assess aspects of competence the program was directly targeted to improve. Typically, this does not mean the very items taught but rather novel items structured similarly to the content addressed in the program. For example, if a program taught word attack, a targeted measure would be decoding of pseudowords. If a program taught comprehension of cause-effect passages, a targeted measure would be answering questions about cause-effect passages structured similarly to those used during intervention, but not including the very passages used for intervention.*

*Broader measures assess aspects of competence that are related to the skills targeted by the program but not directly taught in the program. For example, if a program taught word-level reading skill, a broader measure would be answering questions about passages the student reads.*

*Desired State/Optimal:* The student outcomes measured in the study address at least one of reading domain, with at least one distal/broader measure. For PreK-1 outcomes, PA, vocabulary, or reading comprehension is measured. For Grades 2-6 outcomes, reading comprehension is measured.

*Acceptable:* The student outcomes measured in the study address at least one reading domain using only proximal/targeted measures. Comprehension is one outcome measure but not matched to the desired grade level expectations (e.g., only listening comprehension in Grades 2-6).

*Undesired State:* The student outcomes measured in the study do not address any reading domains OR only measures over- aligned (e.g., tailored to the intervention, relies on materials used in the intervention but not the comparison such as specific reading passages) with the intervention are measured. Neither listening nor reading comprehension is measured.

**h. Treatment Acceptability**

Treatment acceptability refers to an evaluation of the relevance, importance, pragmatics, and/or effectiveness of an intervention for practitioners and students. The evaluation may look at such areas as teacher ownership, teacher appeal, teacher perception of feasibility of implementation, suitability of the intervention to meet student needs, student engagement, or student appeal.

*Desired State/Optimal:* Data are provided related to acceptability of the intervention by teachers and students participating in the study with positive findings.

*Acceptable:* Data are provided related to acceptability of the intervention by teachers or students participating in the study with positive findings.

*Undesired State:* No data on acceptability of the intervention by teacher or students OR negative findings on treatment acceptability.

## **Findings**

**a. Overall Findings (this information is available in reviews on NCII under “Effect Size”)**

Two areas of impact are examined:

- 1) Statistical tests of postintervention differences between study groups on reading outcomes
- 2) Effect size data to represent the magnitude of the relationship between participating in an intervention and the reading outcomes

The effect size is calculated using Hedge's  $g$ , corrected for small-sample bias. First, the adjusted posttest mean for the comparison group is subtracted from the adjusted posttest mean for the intervention group and then divided by the pooled unadjusted posttest standard deviation. This number is multiplied by  $1 - \frac{3}{4N-9}$ .

4N-9.

The adjusted posttest means (i.e. posttests that have been adjusted to correct for any pretest differences between the intervention and comparison groups) will be used to calculate effect sizes. Effect sizes on unadjusted posttest means will only be used/calculated in cases where pretest differences on the measure fall within .25SD and are not statistically significant.

The What Works Clearinghouse guidelines for an ES of .25 or higher considered substantively important is employed in the ratings. The following criteria are applied first to the immediate posttest findings following intervention, and then separately to any long-term findings after the intervention has been completed for 3 or more weeks, or in the case of single case design findings during the generalization/maintenance phase.

*Desired State/Optimal:* The study findings demonstrate statistical significance and an effect size of .25 or higher on one or more reading domains (**see reading domains table on P. 2**) and no statistically significant negative effects. In the case of Intensive Interventions where a single case design is used, the reading data must demonstrate an effect for each case.

*Acceptable:* The study findings demonstrate statistical significance or an effect size of .25 or higher on one or more reading domains and no statistically significant negative effects. In the case of Intensive Interventions where a single case design is used, the reading data must demonstrate at least 3 instances of an effect.

*Undesired State:* The study findings demonstrate no statistical significance or effect sizes of .25 or higher on any reading domains or statistically significant negative effects on one or more reading domains. In the case of Intensive Interventions where a single case design is used, the reading data demonstrate fewer than three instances of an effect.

## Summary of Evidence Across Studies

After completing the ratings on internal validity, external validity, and findings for each unique, primary analysis study of the program submitted, the following summary ratings should be completed taking into account all of the studies submitted.

### a. **Participants (this information is available for each individual study in reviews on NCII under "Participants")**

For this category use the information on the general information page to determine the intended grade levels for the program.

*Desired State/Optimal:* Across the studies submitted, the participants represent all of the intended grade levels for the intervention and the findings are disaggregated for each grade. The participants are clearly described in all studies (e.g., number of students, race/ethnicity, socioeconomic status, disabilities, type of school, etc.). For Classwide Intervention, the participants across studies represent a range of ability levels. For Standardized Treatment Protocols and Intensive Interventions, the participants represent students at-risk for reading difficulties on pre-intervention measures of reading (e.g., below the 30<sup>th</sup> percentile on normed tests, below benchmark on universal screeners).

*Acceptable:* Across the studies submitted, the participants represent some of the intended grade levels for the intervention OR the findings are not disaggregated by grade. Some or all studies provide only partial information regarding the participants. For Classwide Intervention, the participants across studies represent a restricted range with only two groups studied (e.g., high, middle, low performers). For Standardized Treatment Protocols and Intensive Interventions, the participants are identified as at-risk prior to intervention by other criteria (e.g., teacher nominated as struggling, performing in the lowest part of the class).

*Undesired State:* There are no studies that address the intended grade levels for the intervention. Participants are not clearly described in any studies. For Classwide Intervention, the participants across studies represent a restricted range of ability (e.g., only low performers). For Standardized Treatment Protocols and Intensive Interventions, the participants are not identified as at-risk prior to intervention or no information describing how risk status was determined is provided.

#### **b. Extent of Evidence**

*Desired State/Optimal:* Two or more studies with high internal and external validity and positive findings along with no studies showing negative findings. In the case of Intensive Interventions where a single case design was used, five or more studies by three different research teams with high internal and external validity and positive findings.

*Acceptable:* One study with high internal and external validity and positive findings with no studies showing negative findings. In the case of Intensive Interventions where a single case was used, five or more studies with high internal and external validity and positive findings (may all be conducted by 1-2 research teams).

*Undesired State:* No studies with high internal and external validity OR no positive findings. In the case of Intensive Interventions where a single case was used, fewer than five studies with high internal or external validity.

## **Additional Evidence**

### **Additional Evidence:**

Provide a description of any additional evidence of treatment impact or feasibility submitted by the vendor.

### **Subgroup Findings**

Identify whether findings for any subgroups (e.g., at-risk, ELLs, low SES, etc.) are provided separate from the full study sample rated above. If yes, briefly describe the subgroup(s) and findings.

### **Student Outcomes**

List the outcomes that were measured in each study.

## **Resources**

National Center on Intensive Interventions. (2013). *Academic intervention TRC study quality and effect size rating rubrics*.

<http://www.intensiveintervention.org/sites/default/files/pdf/NCIIAcadInterventionRatingRubric2013.pdf>

What Works Clearinghouse. (2014). *Procedures and standards handbook* (version 3.0).

[http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v3\\_0\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf)

---

**B. *Quality  
Review Rubric***

This rubric has all the criteria in a usable review rating form. Generally, there are three levels of ratings: Desired, Acceptable and Undesired. Suggested scores are provided in each rating cell of the rubric in red. The rubric is designed to be used in conjunction with the Reviewer Guide.

---



# Criteria for Classwide, Standardized Treatment Protocol, and Intensive Reading Intervention – General Information

**Intended Grade Levels (mark all that apply):**

PreK 3      PreK 4      PreK 5      K      1      2      3      4      5      6      Above 6

Reading Domains Covered (mark all that apply):		
Print Concepts	Oral Language	<b>Additional Components Covered</b>
Phonological Awareness	Vocabulary	Physical Well-Being and Motor Development
Alphabet Knowledge	Comprehension	Social-Emotional Development
Phonics/Decoding	Spelling	Mathematics and/or Science
Irregular/Sight Words	Writing	Creative Arts
Fluency	Other:	Other:

Recommended Dosage of Program:	
# of weeks of instruction:	N/R
# of minutes per session:	N/R
# of sessions per week:	N/R

Number of Lessons Available:	
Base Materials:	N/A
Extended Materials:	N/A
<b>Placement Assessment Included:</b> Yes      No      N/A	

**Intended Population of Students (mark all that apply):**

General Classroom    Reading/Learning Disabilities    Reading Difficulties    English Language Learners    Language Disabilities    Other: \_\_\_\_\_

**Recommended Implementers (mark all that apply):**

Classroom Teacher    Special Education Teacher    Reading Specialist/Interventionist    Paraprofessional    Peer Tutor    Other: \_\_\_\_\_

**Recommended Grouping Formats (mark all that apply):**

Large Group/Whole Class    Small Group (Group Size \_\_\_\_\_)    Partners (Peer or other)    One-on-One    Other: \_\_\_\_\_

**Parent/Home Connection Strategies/Materials Included:**      yes      no

**Number of Studies Submitted:** \_\_\_\_\_ **Number of Peer-reviewed Studies Submitted:** \_\_\_\_\_

Costs		
Material Costs (per pupil):	N/A	N/R
Training Costs:	N/A	N/R
Hours of training recommended:	N/A	N/R
Additional Costs: Extended Materials or Additional Training Required:	N/A	N/R

# Evidence for Program Effectiveness

Evidence				
	Internal Validity	Desired State/ Optimal	Acceptable	Undesired State
a. Study Design	Research design	Random assignment of participants to study conditions AND no significant overall or differential attrition  <b>[40 points]</b>	Random assignment of participants with evidence of significant overall or differential attrition OR quasi-experimental study with study conditions matched on several pretreatment variables OR single case design with at least 3 replications of effect and at least 3 data points per phase of design  <b>[20 points]</b>	Other study designs (e.g., single group, case study, etc.) OR quasi-experimental study without adequate matching OR single case design with less than 3 replications of effects or less than 3 data points per phase of design  <b>[0 points]</b>
	Comparison Group	Intervention and comparison groups are not statistically significantly different at pretreatment OR baseline data are stable for all cases prior to intervention implementation  <b>[40 points]</b>	Intervention and comparison groups are statistically significantly different at pretreatment AND are within .25 SD difference at pretreatment, AND outcomes are adjusted for pretreatment differences in analysis OR baseline data are stable for the majority of cases (must be more than 3 cases)  <b>[20 points]</b>	Intervention and comparison groups are above .25 SD difference at pretreatment OR baseline data are not stable  <b>[0 points]</b>
b.	Group/Person Conducting Study	Study conducted by independent evaluators without conflict of interest in the intervention (e.g., developers of the intervention financial interest in the intervention)  <b>[20 points]</b>	Study conducted by those with possible conflict of interest (e.g., same developer but not with financial interest in the intervention)  <b>[10 points]</b>	Study conducted by those with a conflict of interest in the intervention  <b>[0 points]</b>
c.	Developer of Assessment	The majority of assessments were developed by someone other than the authors/developers of the intervention  <b>[10 points]</b>	Less than half of the assessment(s) were developed by the vendor but not the author/developer of the intervention  <b>[5 points]</b>	All assessment(s) were developed by authors/developers of the intervention  <b>[0 points]</b>

d. Technical Adequacy	Technical Adequacy of Measures Used to Determine ES or Evidence of Improvement: Reliability	All reliability coefficients $\geq .70$ AND most $\geq .80$ <b>[10 points]</b>	All reliability coefficients $\geq .70$ <b>[5 points]</b>	Most reliability coefficients $< .70$ OR reliability coefficients not provided for most measures <b>[0 points]</b>
	Technical Adequacy of Measures Used to Determine ES or Evidence of Improvement: Validity	All validity coefficients $\geq .30$ AND most $\geq .50$ <b>[10 points]</b>	All validity coefficients $\geq .30$ <b>[5 points]</b>	Most validity coefficients $< .30$ OR validity coefficients not provided for most measures <b>[0 points]</b>
e.	Data Collection	Outcome data collected by assessors blind to study condition AND reliability/intercoder agreement $\geq .90$ OR more than assessor for single case design with adequate reliability/intercoder agreement $\geq .90$ <b>[20 points]</b>	Outcome data collected by assessors blind to part of the study condition AND reliability/intercoder agreement $\geq .80$ <b>[10 points]</b>	Data collectors not blind to study condition OR inadequate reliability/intercoder agreement $< .80$ OR single assessor for single case design <b>[0 points]</b>
f.	Data Analyses	Unit of assignment matches unit of analysis and/or clustering/nesting of students is taken into account AND $p$ values for statistical significance adjusted for multiple comparisons OR single case design is appropriate for measuring the intervention effect <b>[40 points]</b>	Analyses do not take into account clustering/nesting OR $p$ values for statistical significance are not adjusted for multiple comparisons <b>[20 points]</b>	Analyses do not match design of the study <b>[0 points]</b>
g.	Evidence of Confounding Factors (e.g., differences between study groups beyond the intervention of interest)	No evidence of confounding factors that could have interfered with results <b>[10 points]</b>	<b>[5 points]</b>	Evidence of confounding factors that could have interfered with results <b>[0 points]</b>

	External Validity	Desired State/ Optimal	Acceptable	Undesired State
a.	Group/Person Implementing Intervention	Intervention implemented in school setting by school personnel <b>[25 points]</b>	Intervention implemented in school setting with specially trained personnel outside of school personnel (e.g., researcher) <b>[15 points]</b>	Intervention implemented outside of school setting (e.g., clinic) by specially trained personnel <b>[0 points]</b>
b.	Dosage: Session Time and Frequency	Session time and frequency in study were aligned with recommendations for the intervention <b>[10 points]</b>	<b>[5 points]</b>	Session time and frequency in the study were not aligned with recommendations for the intervention <b>[0 points]</b>
c.	Fidelity of Implementation	Intervention implementation was generally high (75% or more) with detailed data on implementation AND adequate intercoder agreement $\geq .90$ <b>[75 points]</b>	Intervention implementation was generally high (75% or more), but either limited data or indirect data (e.g., teacher reports) are provided OR intercoder agreement $\geq .80$ <b>[35 points]</b>	No data related to the implementation of the intervention during the study OR intervention was not implemented as intended OR intercoder agreement is $< .80$ <b>[0 points]</b>
d.	Reading Domains Addressed	Classwide/STP: Three or more foundational reading domains are addressed in the intervention  Intensive (any domain) <b>[50 points]</b>	Classwide/STP: At least two foundational reading domains are addressed in the intervention  <b>[25 points]</b>	Classwide/STP: Fewer than two foundational domains are addressed in the intervention  Intensive (no domains) <b>[0 points]</b>
e.	Grouping Format	Grouping format(s) in study aligned with recommended grouping for intervention <b>[10 points]</b>	<b>[5 points]</b>	Grouping format(s) in study did NOT align with recommended grouping for intervention <b>[5 points]</b>
f.	Student Outcomes Measured: Distal/Broad	Student outcomes address at least one reading domain, with at least one distal/broader measure <b>[10 points]</b>	Student outcomes address at least one reading domain with only proximal/targeted measures of the intervention <b>[5 points]</b>	Student outcomes do not address any reading domains OR only measures over-aligned with the intervention are measured <b>[0 points]</b>
	Student Outcomes Measured: Comprehension	Comprehension is measured (PA or Voc PreK - 1, and reading comprehension in 2 <sup>nd</sup> grade and up) <b>[10 points]</b>	Comprehension is measured but not matched with grade (e.g., listening comprehension only in 2 <sup>nd</sup> grade and up) <b>[5 points]</b>	Comprehension is not measured (PA or Voc PreK - 1, and reading comprehension in 2 <sup>nd</sup> grade and up) <b>[0 points]</b>

g.	Treatment Acceptability (e.g., teacher ownership, teacher appeal, feasibility, student engagement, student appeal)	Data are provided related to acceptability of the intervention by teachers and students participating in the study with positive findings <b>[10 points]</b>	Data are provided related to acceptability of the intervention by teachers or students participating in the study with positive findings <b>[5 points]</b>	No data on acceptability of the intervention by teacher or students OR negative findings on treatment acceptability <b>[0 points]</b>
	<b>Findings</b>	<b>Desired State/ Optimal</b>		<b>Undesired State</b>
a.	Overall Findings	Statistical significance AND effect size of .25 or higher on one or more reading domains AND no statistically significant negative effects OR in single case design the reading data demonstrate an effect for each case <b>[80 points]</b>	Statistical significance OR effect size of .25 or higher on one or more reading domains AND no statistically significant negative effects OR in single case design the reading data demonstrate at least three instances of an effect <b>[40 points]</b>	No statistical significance or effect sizes higher than .25 on any reading domains OR statistically significant negative effects one or more reading domains OR in single case design the reading data demonstrate fewer than three instances of an effect <b>[0 points]</b>
b.	Long-Term Findings	Statistical Significance AND effect size of .25 or higher on one or more reading domains OR in single case design the reading data demonstrate an effect for each case in maintenance/generalization phases <b>[20 points]</b>	Statistical significance OR effect size of .25 or higher on one or more reading domains OR in single case design the reading data demonstrate an effect for at least three cases of an effect in maintenance/generalization phases <b>[10 points]</b>	No statistical significance or effect sizes higher than .25 on any reading domains OR in single case design the reading data demonstrate fewer than three instances of an effect in maintenance/generalization phases <b>[0 points]</b>

## Summary of Evidence Across Studies

		Desired State/ Optimal		Undesired State
a.	Participants	Participants across the studies are in all of the intended grades for the intervention and the findings are disaggregated for each grade <b>[8 points]</b>	Participants across the studies represent some of the intended grades for the intervention OR the findings are not disaggregated by grade <b>[4 points]</b>	No studies address the intended grades for the intervention <b>[0 points]</b>
		Participants are clearly described in all studies (e.g., geographical, race/ethnicity, SES, student population, type of school (dual language, public, magnet) <b>[6 points]</b>	Some or all studies provide only partial information regarding participants (e.g., geographical but no SES) <b>[3 points]</b>	Participants are NOT clearly described in any studies (e.g., geographical, race/ethnicity, SES, student population, type of school: dual language, public, magnet) <b>[0 points]</b>
		Classwide: Participants represent a range of ability levels STP/Intensive: Participants at-risk on pre-intervention measures of reading (e.g., below 30 <sup>th</sup> percentile on normed tests, below benchmark on universal screeners) <b>[6 points]</b>	Classwide: Participants represent a restricted range with only two groups studied (e.g., high, middle, low performers) STP/Intensive: Participants identified as at-risk prior to intervention by other criteria (e.g., teacher nominated as struggling, performing in the lowest part of class) <b>[3 points]</b>	Classwide: Participants represent a restricted range of ability (e.g., only low performers) STP/Intensive: Participants not identified as at-risk prior to intervention OR no information describing how risk was determined <b>[0 points]</b>
b.	Extent of Evidence	Two or more studies with high internal and external validity with positive findings with no studies showing negative findings OR in single case design five or more studies by three different research teams with high internal and external validity with positive findings <b>[80 points]</b>	One study with high internal and external validity with positive findings with no studies showing negative findings OR in single case design five or more studies with high internal and external validity with positive findings <b>[40 points]</b>	No studies with high internal and external validity OR no positive findings OR in single case design fewer than five studies with high internal or external validity <b>[0 points]</b>

---

## C. Additional Rubrics [RFP Review]

There may be additional review areas needed – therefore two additional rubrics are provided as possible support to review teams. The same review procedures should be followed, regardless of the rubric used [described in 1. *Review Guide and Quality Review Rubric*]:

1. **Standards Alignment Criteria Rubric.** This rubric was developed by Dr. Andy Porter. Suggested scores are provided in each rating cell of the rubric in red. The rubric includes the following areas:
    - a Criterion A: Standard Identification
    - b Criterion B: Amount of Alignment
    - c Criterion C: Replicability
    - d Criterion D: Content Definition & Specificity
    - e Criterion E: Inter-rater Reliability
    - f Criterion F: Conductor of Alignment
    - g Criterion G: Misalignments
  2. **Teacher Usability Rubric.** This rubric was adapted from a rubric used by the Iowa Reading Research Center in 2014 in their parent resource reviews. The rubric includes the following areas:
    - a Accessibility
    - b Credibility
    - c Content
    - d Bias
    - e Rater Use/Grade
-

# Standards Alignment Criteria

Used in RFP Review only

## Criterion A: Standard Identification

There is a clear and published statement of what parts of the Common Core State Standards (CCSS) are addressed in the materials/program and how well-aligned the materials are to those standards.

The statement should address the subject, grade level(s), standards, and specific content in those standards that are covered. Perfect alignment indicates a one-to-one correspondence between content in the target standards and content in the materials/program.

The materials/program can be more or less focused as to grade level(s) and CCSS standards. This criterion is met to the extent (a) the target standards are clearly identified, the range of standards addressed is broad, and the alignment is high.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
A.	Standard Identification	The content to be taught and learned is clearly identified with CCSS standards, the breadth of standards covered is instructionally important (e.g. all standards for a grade level), and the content alignment for the standards is claimed to be excellent. <b>[10 points]</b>	The relationship between the program/materials and CCSS standards is addressed but either the language is not precise, the breadth of standards covered is modest, and/or the alignment is modest (e.g. important content in the target standards are not addressed). <b>[5 points]</b>	There is no description of what CCSS standards are addressed.  <b>[0 points]</b>



**Criterion B: Amount of Alignment**

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
B.	Amount of Alignment	Alignment exists in degree. The Webb procedure, based on four criterion, assesses alignment between an assessment and content standards. A minimum acceptable amount of alignment is set for each criterion. The SEC has a single index of alignment and can be used for judging alignment between any two statements of content, (e.g. standards, assessments, curriculum materials, and instruction). Alignment between test and standards is excellent at 0.5. Alignment between a comprehensive set of materials and standards should be higher than 0.5. <b>[10 points]</b>	The degree of alignment is reported and is reasonable but below minimum threshold on one or more dimensions.  <b>[5 points]</b>	Degree of alignment is either not reported or reported as being unacceptably low.  <b>[0 points]</b>

### **Criterion C: Replicability**

A replicable methodology is used to measure and describe alignment.

The methodology is sufficiently clear and objective that the claims for alignment could be reproduced by an independent effort.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
C.	Replicability	The methodology used to measure degree and nature of alignment is well-specified, objective, and could be replicated.  <b>[20 points]</b>	The methodology is well-described in terms of expertise of individuals judging alignment. Judges work independently and the average across judges is taken. Consensus clinical judgment is a useful approach but yields unknown reliability.  <b>[10 points]</b>	One individual of unknown expertise judges the alignment.  <b>[0 points]</b>

### **Criterion D: Content Definition & Specificity**

When judging alignment between content of the CCSS and content of the materials/program, content is clearly defined and at a level of precision useful for capturing the match in content at a level of detail likely to capture opportunity to learn (OTL) that is necessary for mastery of the targeted standards.

There are various approaches to defining content among methodologies for assessing content alignment. Research has found that content needs to be distinguished at least at the intersection of topics and cognitive demand. Topics defined the types of content (e.g. linear equations, self-correction strategies in reading comprehension). Cognitive demand, distinguishes what about the topic students are to know or be able to do (e.g. solve a linear equation, versus use a linear equation to solve a story/problem).

There is no clear agreement on the level of detail most useful but there should be evidence provided that the number of topics and number of levels of cognitive demand are sufficient to describe OTL precisely enough so that OTL is predictive of student achievement.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
D.	Content Definition & Specificity	Alignment is judged at the intersection of topics and cognitive demand and research is cited demonstrating that the level of detail (grain size) is appropriate for making important distinctions in OTL.  <b>[30 points]</b>	Content is defined but alignment is judged and reported separating for topics and cognitive demand of only very large chunks of topics or crude distinctions among levels of cognitive demand are addressed.  <b>[15 points]</b>	Content is not defined.  <b>[0 points]</b>

### **Criterion E: Inter-rater Reliability**

There is sufficient reliability of the content judgments being made to assess the amount and nature of alignment between the CCSS and the materials/program.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
E.	Inter-rater Reliability	When measuring alignment, multiple experts are used and the agreement among experts is investigated, reported, and high (e.g. reliability of the average judgment .65 or above). <b>[10 points]</b>	Only a single expert is used in judging alignment or multiple experts are used but interrater agreement is not reported. <b>[5 points]</b>	Reliability is not considered. <b>[0 points]</b>

### **Criterion F: Conductor of Alignment Study**

An independent third party investigates and reports on the content alignment of the materials/program to the CCSS.

The producers, authors, vendors of the materials/program have a vested interest in their product being aligned to the CCSS. An independent third party assessment is less likely to let bias toward positive findings of alignment creep into the reported results.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
F.	Conductor of Alignment Study	The alignment investigation and reporting was done by an independent third party. <b>[15 points]</b>	The alignment investigation and reporting were done by the author who is not the vendor and there is no profit incentive. <b>[8 points]</b>	The alignment investigation and reporting was done by the for-profit vendor. <b>[0 points]</b>

### **Criterion G: Misalignments**

The content in the target standards not covered in the materials/program as well as the content in the materials/program not covered in the target CCSS standards is clearly described.

Standards Alignment				
	Criterion	Desired State/ Optimal	Acceptable State	Undesired State
G.	Misalignments	Beyond stating the degree of alignment, additional content not aligned as well as relevant standards from the CCSS not addressed are clearly explained.  <b>[5 points]</b>	A general description of content in the targeted standards not covered in the materials/ program is provided but no description is provided for content in the materials/program not found in the targeted standards (or vice versa).  <b>[3 points]</b>	There is no description of either the content in the targeted standards not covered in the materials/ program nor the content in the materials/ program not in the targeted standards.  <b>[0 points]</b>

# Teacher Usability Rubric

Used in RFP Review only

## TEACHER USABILITY: ACCESSIBILITY, CREDIBILITY, MATERIALS AND FREE OF RACE/GENDER BIAS

Intervention Number:

Reviewer:

ACCESSIBILITY: This category refers to the 'user-friendliness' of the content or resource.				SCORE
3	2	1	DNA	
<b>Appearance.</b> For video, this refers to the quality of images and audio. For apps and websites it refers to the layout and labelling of content. For text resources it refers to formatting of the document (e.g., bullet points, font size and graphics make the content easy to use). This refers to the entire page.				
This rating indicates that ALL of the contents are of high quality. Audio, video, layout, etc. are attractive and draw the user's attention.	This rating indicates that SOME of the contents are of high quality. For example, the video may be clear but the audio is not. The text content is attractively laid out but difficult to read.	This rating indicates that NONE TO LITTLE of the material is of acceptable quality. For example, a web page is cluttered; it is difficult to see how to advance to other areas in the website, etc.		
<b>Does not require registration or fee.</b> Is the user required to provide personal information or pay a fee to access the content?				
The user can access the content without revealing personal information or paying an access fee. There is no requirement to create a username or password. Students can access activities and/or materials without mandatory access to a laptop or materials not normally available within the classroom.	The user can access SOME of the content BUT only after providing personal information OR paying an access fee. For example, if the user wants to see all of the content, they must create a username and password.	The user must provide personal information or pay an access fee in order to engage with ANY of the content. Students MUST have access to a laptop or other expensive materials/equipment not normally available within the classroom.		



<b>Content is presented in simple, clear language.</b> Content contains words/jargon that all teachers understand and written and spoken in simple language (free of heavy use of acronyms and multiple words that will have to be defined with a dictionary) that most teachers could understand. If the content provides suggestions for activities, step-by-step instructions are clear and easy to follow.				
This rating indicates that ALL of the contents meet the criteria. ALL of the language is easy to follow and understand. For example, there are no acronyms, all words are understandable, and instructions are clear.	This rating indicates that SOME of the contents meet the criteria. Most of the content is written in simple, clear language. For example, some acronyms may be used, but they are explained.	This rating indicates that NONE TO LITTLE of the material meets the criteria. Most of the language contains jargon or is confusing. For example, multiple acronyms are used and many words are present that will need to be defined using a dictionary.		
<b>Text content is accessible.</b> Is the text of the content downloadable? Can it be saved or exported? Can it be printed? <b>NOTE: This is not applicable to video resources or to apps.</b>				
The text content fully meets the criteria. Users can save the content to access it offline and in hard copy format.	The contents partially meet the criteria. For example, users can bookmark the content but not print it.	The content is only available within the digital platform. Users cannot bookmark, copy and paste, download or save the content.		
<b>TOTAL SCORE: ACCESSIBILITY</b>			<b>/15</b>	
<b>Notes:</b>				
<b>CREDIBILITY:</b> This category refers to the trustworthiness or reputability of the resources.				<b>SCORE</b>
<b>Content is objective.</b> Content does not contain ads. If content contains ads or is a commercial product, the content is focused more on education than on sales. An objective source would not 'push' a particular product or point of view. The majority of the screen should be devoted to content, not ads.				
The resource appears intended to educate the user, not sell a product. Language used to describe the resource avoids salesmanship. Ads are small in size and are differentiated from the content.	Advertisements are present but they are less prominent than the educational content. The goal appears more to educate than to sell.	Advertisements dominate the page and make unrealistic claims or if an article recommends a particular instructional tool to purchase.		

<b>Content is current (may not apply to all resources).</b> Copyrights are one indicator of how recently the content was developed or revised. For text resources indicators include date of publication. For web sites, the date the site was last updated is important.				
The content has been updated in the past year or has a copyright within the past 5 years.	Copyright, date of publication or updated indicate that the content is 6-10 years old.	Copyright, date of publication or updated indicate that the content is 10+ years old or the information cannot be found.		
<b>TOTAL SCORE: CREDIBILITY</b>			<b>/6</b>	
<b>Notes:</b>				
<b>CONTENT:</b> This refers to the contents' potential for supporting ease of teacher use in the classroom.				<b>SCORE</b>
<b>TEACHER - Suggests activities that teachers can use during activities.</b> Activities do not require purchasing specific materials, space or other resources. Suggested materials are easily found or accessible.				
This rating indicates that ALL of the contents involve activities and materials that are commonly available.	This rating indicates that SOME of the contents involve activities and commonly available materials. For example, it might ask for easy access to laptops or specific student stimuli that would be outside of the normal classroom.	NONE TO LITTLE of the contents involve activities and commonly available materials. This might require an additional purchase of a set of materials beyond the common set of materials.		
<b>STUDENT - Suggested activities are engaging and interactive.</b> The child is speaking, reading, listening, pointing, clicking or some response from child involved.				
The suggested activities involve a high degree of engagement from the child. The child is required to elicit responses consistently throughout the activity.	The suggested activities involve some level of engagement from the child. The child is passive or responding a few times.	NONE of the suggested activities require engagement from the child. No response is needed from the child		

<b>TEACHER - Intention or purpose of engaging with the resource is clear.</b> The literacy skills the child will develop through engaging with the resource are clearly identified and explained. Skills might include: phonemic awareness, phonics, fluency, comprehension, vocabulary, writing, and spelling.				
The literacy skills targeted by the resource are explicitly identified and explained. For example, the resource explains that vocabulary development is a critical skill in order to achieve reading proficiency.	The literacy skills targeted by the resource are identified but NOT explained. For example, the resource explains that it can support children's vocabulary development but does NOT give reason behind the importance of vocabulary development.	The literacy skills targeted by the resource are NEITHER explicitly identified nor explained.		
<b>TOTAL SCORE: CONTENT</b>			<b>/9</b>	
<b>Notes:</b>				
<b>FREE OF RACE AND/OR GENDER BIAS</b>				
<b>Materials do not stereotype across any group (e.g., socio-economic status, gender, race, religion, homosexuality, ability, familial status, ancestry).</b>				
Strongly Agree - 3	Agree - 2	Disagree - 1	Strongly Disagree 0	
<b>Materials are devoid of discriminatory language.</b>				
Strongly Agree - 3	Agree - 2	Disagree - 1	Strong Disagree - 0	
<b>TOTAL SCORE: FREE OF RACE/GENDER BIAS</b>			<b>/6</b>	
<b>Notes:</b>				



TOTAL SCORE			
Accessibility	Credibility	Content	Bias
		TOTAL	
1. I would recommend this resource to a teacher Agree = A      Neutral = B      Disagree = C		RECOMMENDATION 1	
2. I would use this resource. Agree = A      Neutral = B      Disagree = C		RECOMMENDATION 2	

